Model of Transformation of the Alphabet of the Encoded Data as a Tool to Provide the Necessary Level of Video Image Quality in Aeromonitoring Systems

Serhii Khmelevskiy, Ivan Tupitsya, Maksim Parkhomenko and Yan Borovensky

Ivan Kozhedub Kharkiv National Air Force University, Sumskaya Str. 77/79, Kharkiv, 61000, Ukraine

Abstract

The drawbacks of the existing algorithms for coding data of video information resources in the aero monitoring system are analyzed. The main ones are: the complexity of the algorithmic implementation; loss of informative data that determine the semantic component (carry semantic load) of the information resource. To solve these problems, a model for transforming the alphabet of encoded data is being developed. The essence of the developed model is to determine and take into account the significance of the elements of the encoded data. The significance of message elements means taking into account the probabilistic distribution of elements in the message and the structural features of the color model to which the initial data is presented. Experimental studies are being carried out to confirm the adequacy of the developed model. The effectiveness of the developed model is assessed from the standpoint of ensuring the reduction of the power of the alphabet of the encoded data, provided that the high quality of video images is ensured. The directions of further research are determined, which involve the synthesis of the developed model with technologies for coding data information resources without losses.

Keywords [1]

transformation, information resource data, redundancy, significance, coding, aeromonitoring

1. Introduction

In the system of information support of departmental bodies, a rather important role is assigned to the use of the aero segment [1-3]. This is due to the fact that obtaining information in real time allows for both timely detection and prompt response to crisis situations by organizing coordinated interaction of the relevant departmental bodies.

Accordingly, the requirements for the video information resource are increasing. The main ones among them [4-8] are: a compact representation of the encoded data, a high level of quality of reconstructed images in conditions of limited bandwidth of data transmission channels. For this purpose, image coding algorithms based on conceptual approaches, which are implemented on the basis of the JPEG platform [9-17], are quite actively used. However, it should be noted that algorithms of the JPEG family have a number of significant drawbacks [12-27]. The main disadvantages include the following: the complexity of the algorithmic implementation; loss of informative data (key information) that determine the semantic component (carry semantic load) of the information resource. Therefore, the question of finding new approaches to the compact representation of encoded data in terms of ensuring a high level of quality becomes relevant.

The analysis of the latest scientific research indicates that the existing compression technologies are built on the principle of multilevel processing (they assume the implementation of step-by-step data processing) [28-34]. This means that the original alphabet is used only at the initial stage of video

Information Technology and Implementation (IT&I-2021), December 01–03, 2021, Kyiv, Ukraine

EMAIL: sserg1978@ukr.net (S. Khmelevskiy); ivan20081982@gmail.com (I. Tupitsya); maxpar76@gmail.com (M. Parkhomenko); borovenskii2001@gmail.com (Ya.A. Borovensky)

ORCID: 0000-0001-6216-3006 (S. Khmelevskiy); 0000-0001-6806-4914 (I. Tupitsya); 0000-0001-6062-7743 (M. Parkhomenko); 0000-0002-8747-4203 (Ya.A. Borovensky)

data processing. All subsequent stages (preparation for encoding and encoding, respectively) are associated with the processing of intermediate data. And accordingly, an increase in the number of processing stages implies an increase in the complexity of the algorithmic implementation (which increases the time for data processing) and a decrease in the quality of the restored images.

Therefore, it is proposed to use an approach to ensure the required level of video image quality under conditions of limited bandwidth of data transmission channels, the essence of which is to transform the original alphabet of the encoded data in the direction of reducing psycho-visual redundancy by determining the significance of the encoded data elements.

Thus, the purpose of the article is to develop a model for transforming the alphabet of encoded data to ensure the required level of video image quality in conditions of limited bandwidth of data transmission channels in aero monitoring systems.

2. Development of a transformation model for the alphabet of encoded data to ensure the required level of video image quality in aeromonitoring systems

It is proposed to develop a model for transforming the alphabet of the encoded data, which will take into account the significance of the encoded elements according to some quantitative criterion - the sign of the element's significance. So for the message, which is given by the following expression:

$$X(n) = \{x_1; ...; x_i; ...; x_n\}, \ i = \overline{1, n}, \quad (1)$$

where x_i - i - th message element $X(n)$, $i = \overline{1, n}$, message alphabet $X(n)$ will look like this:

$$X(m) = \{x_1; ...; x_i; ...; x_m\}, \ i = \overline{1, m}, \quad (2)$$

where m - the number of elements in the alphabet $X(m)$ messages $X(n)$, $i = \overline{1, m}$.

Taking into account the fact that at the initial stage the encoded data is represented using the RGB model, that is, they are elements of an RGB cube, the elements of the encoded data can be represented as follows:

$$x_i = x_i(x_{iR}, x_{iG}, x_{iB}), \quad (3)$$

where x_{iR} - the component *red* of the element x_i; x_{iG} - the component *green* element x_i; x_{iB} - the component *blue* of the element x_i.

It is proposed to transform the alphabet of the encoded data in several stages (Fig. 1). At the first stage, the elements are determined x_i messages $X(n)$, which do not have a significant impact on the semantic component of the message, that is, psycho-visual redundancy is eliminated. At this stage, the probability distribution law is determined $P(x_i)$ appearance of elements x_i in the message $X(n)$.

1. At the next stage, the significance of the elements is determined x_i messages $X(n)$ quantitatively. To this end, the element x_i messages $X(n)$ represented as the center of a cube in the RGB color space with the size κ^3, i.e.:

$$V_{RGB} = \kappa^3, \quad (4)$$

where V_{RGB} - the volume of the cube in the RGB color model;
κ - the length of the side of the cube, in pixels.

To assess the significance of elements x_i messages $X(n)$ it is proposed to use a quantitative feature, which is given by the following expression:

$$S(x_i) = \frac{P(x_i)}{\sum_{j=1}^{\alpha} \frac{P(x_i)}{D(x_i, x_{ij})^{\beta}}}, \quad (5)$$

where $S(x_i)$ - the coefficient of significance of the element x_i messages $X(n)$;
$P(x_i)$ - the probability of the item appearing x_i in the message $X(n)$;
x_i - message element $X(n)$, which is the center of the generated cube in the RGB color model;
x_{ij} - elements of the formed color cube in the RGB model, $x_{ij} = x_{ij}(x_{ij_R}, x_{ij_G}, x_{ij_B})$, $j = \overline{1,\alpha}$, $\alpha = 3(\kappa - 1)$;
$D(x_i, x_{ij})$ - distance between element x_i (the center of the formed cube) and elements x_{ij} the generated RGB cube.

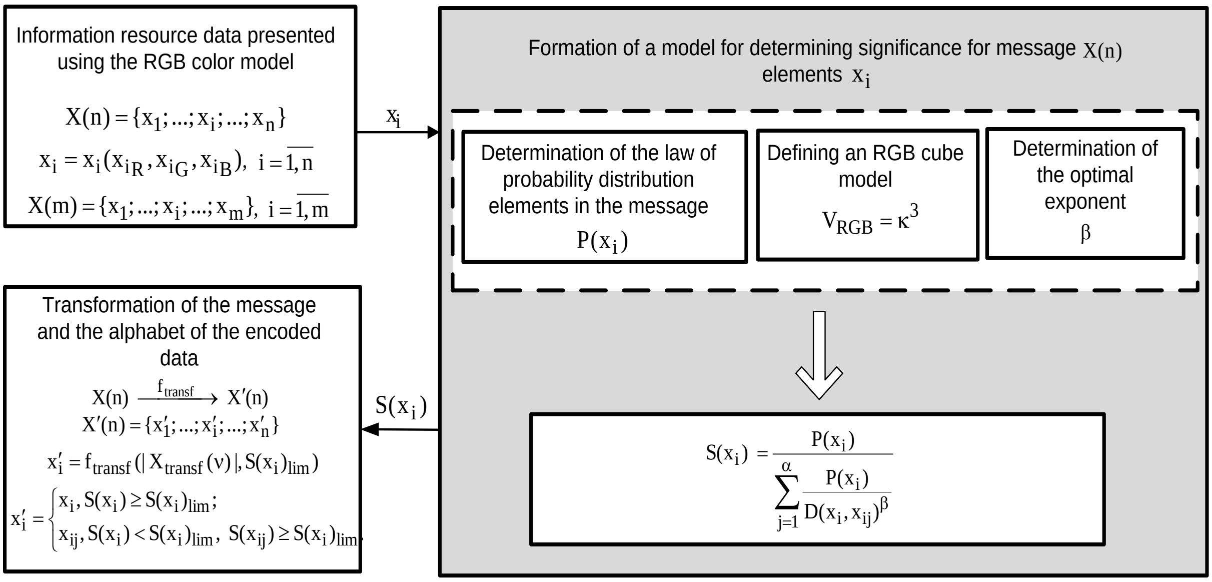

Figure 1: Structural and functional diagram of the developed model of transformation of the alphabet of encoded data

In turn, the distance $D(x_i, x_{ij})$ between element x_i messages $X(n)$ and element x_{ij} An RGB color cube is defined by the following expression:

$$D(x_i, x_{ij}) = \sqrt{(x_{iR} - x_{ij_R})^2 + (x_{iG} - x_{ij_G})^2 + (x_{iB} - x_{ij_B})^2}\ . \qquad (6)$$

2. The third stage is the ranking of the data array (coefficients of significance $S(x_i)$ elements x_i messages $X(n)$) from maximum to minimum value.

3. At the fourth stage, the significance threshold is determined $S(x_i)_{lim}$ elements x_i messages $X(n)$ (image pixels). This indicator is determined by the power of the alphabet of the reconstructed message. Thus, taking into account the transformation of the alphabet $X(m)$ initial data using the proposed approach message $X(n)$ will take the following form:

$$X(n) \xrightarrow{f_{transf}} X'(n),\ X'(n) = \{x'_1; ...; x'_i; ...; x'_n\}, \qquad (7)$$

where $f_{transf}(|X_{transf}(\nu)|, S(x_i)_{lim})$ - function of forming a transformed message taking into account the threshold value of the significance coefficient $S(x_i)_{lim}$ elements and power of the alphabet of the transformed message;

x'_i - i - th element transformed message $X'(n)$, $i = \overline{1,n}$;

$X_{transf}(\nu)$ - transformed message alphabet $X'(n)$, which is formed taking into account the significance indicator;

$|X_{transf}(\nu)|$ - the power of the alphabet of the transformed message $X'(n)$.

This uses the significance information (values of the coefficient of significance $S(x_i)$) elements messages $X(n)$ to form a transformed message $X'(n)$. Given this information, based on the function $f_{transf}(|X_{transf}(\nu)|, S(x_i)_{lim})$ items are ranked x'_i transformed message $X'(n)$ at a given level power $|X_{transf}(\nu)|$ alphabet $X_{transf}(\nu)$ transformed message $X'(n)$. This is given by the following expression:

$$x'_i = f_{transf}(|X_{transf}(\nu)|, S(x_i)_{lim}). \tag{8}$$

It should be noted that it is from the power $|X_{transf}(\nu)|$ alphabet $X_{transf}(\nu)$ transformed messages $X'(n)$ the degree of compression and, accordingly, the degree of distortion of the image elements depends. Thus, at the final stage, the formation of the alphabet takes place $X_{transf}(\nu)$ transformed messages $X'(n)$, the essence of which is to replace the elements x_i original message $X(n)$ for which the significance coefficient $S(x_i)$ matters less than the threshold $S(x_i)_{lim}$, that is:

$$S(x_i) < S(x_i)_{lim},$$

elements of the alphabet $X_{transf}(m)$ for which the following condition is satisfied:

$$S(x'_i) \geq S(x_i)_{lim}.$$

Thus, the process of forming a transformed message $X'(n)$ is given by the following expression system:

$$x'_i = \begin{cases} x_i, S(x_i) \geq S(x_i)_{lim}; \\ x_{ij}, S(x_i) < S(x_i)_{lim},\ S(x_{ij}) \geq S(x_i)_{lim}. \end{cases} \tag{9}$$

In turn, the purpose of transforming the alphabet $X(m)$ the encoded data is power reduction $|X(m)|$ subject to the provision of high quality video images (in terms of using compression algorithms for coding data), i.e. the following condition must be met:

$$|X_{transf}(\nu)| < |X(m)|. \tag{10}$$

3. Evaluation of the effectiveness of the developed model of transformation of the alphabet of encoded data from the standpoint of ensuring the required level of video image quality in aeromonitoring systems

Analysis of transformations of the nature of the probability distribution law for the appearance of message elements as a result of clustering based on the number of series of units.

In order to determine the optimal values of quantitative indicators (RGB cube model - length κ sides of a cube, exponent β), which are used in the developed model of transformation of the alphabet $X(m)$ coded data, it is proposed to conduct a number of experimental studies. To determine the effectiveness of the developed model for transforming the encoded data, a number of experiments were carried out. Highly saturated images were used as initial data; an example of one of them is shown in Fig. 2. The results of experimental studies to determine the optimal values of quantitative indicators that are used in the developed model of the transformation of the alphabet $X(m)$ of the encoded data are shown in Figs. 3 - 4. Analysis of the results of experimental studies to determine the optimal values of quantitative indicators that are used in the developed model of transformation of the alphabet $X(m)$ coded, which are presented in Fig. 3-4, indicate that:

1. The optimal value for the quantitative indicator of the RGB cube model is length κ side of the cube is as follows (Fig. 3):

$$\kappa = 9. \tag{11}$$

This is due to the fact that for a given value of the length κ side of the cube, the minimum standard deviation is provided, which has the following value (Fig. 3):

$$CKO = CKO_{min} \approx 0,7\% .$$

It should be noted that this RGB cube model (with $\kappa = 9$) is selected taking into account the assessment of the complexity of the algorithmic implementation of the developed model for transforming the encoded data (Fig. 4).

2. Analysis of the dependence of the standard deviation on the length κ sides of a cube for different values of the exponent β (Fig. 4) indicates that for the developed model of transformation of the alphabet of encoded data with the optimal value of the indicator β is the following:

$$\beta \approx 3,5. \quad (12)$$

At $\beta = 3,5$ the standard deviation is provided, which has the following value (Fig. 4):

$$CKO \approx 0,7\% .$$

Taking into account the results of the experimental studies (data from expressions (11) - (12)), expression (5) takes the following form:

$$S(x_i) = \frac{P(x_i)}{\sum_{j=1}^{24} \frac{P(x_i)}{D(x_i, x_{ij})^{3,5}}} . \qquad \frac{|X(m)|}{|X_{transf}(\nu)|} = 9,55. \quad (13)$$

Further, it is proposed to evaluate the effectiveness of the developed model from the standpoint of ensuring power reduction $|X(m)|$ alphabet $X(m)$ of the encoded data, provided that the high quality of video images is ensured, that is, the condition specified by expression (10) is met. So for the tested image, which was used in the course of experimental studies, the result of applying the developed model for transforming the alphabet of the encoded data is a decrease in the power of the alphabet by approximately 10 times (Fig. 5):

Figure 2: Test highly saturated image

Thus, we can conclude that the proposed model for transforming the alphabet of the encoded data allows to ensure not only the high quality of reconstructed images, but also to create more favorable conditions for the further use of lossless compression coding technologies in order to compactly represent the data of the information resource. Therefore, the goal of further research is to synthesize the developed model for transforming the alphabet with technologies for compact representation of encoded data without loss to increase the efficiency of delivery of data information resources while ensuring the required level of quality.

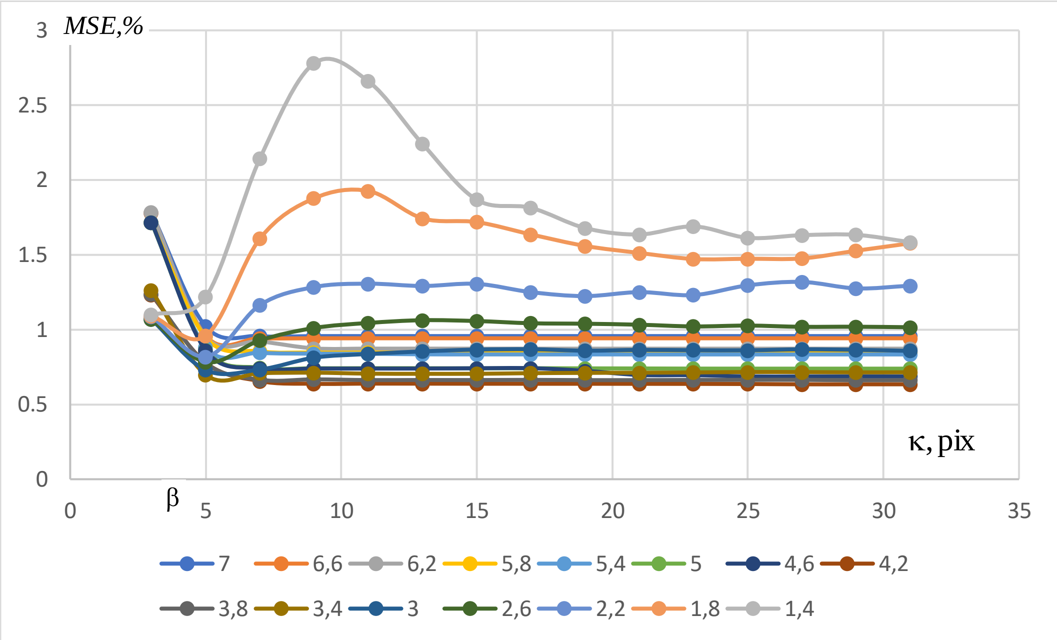

Figure 3: Diagram the dependence of the standard deviation on the length κ sides of a cube for different values of the exponent β for the developed video sequence alphabet transformation models

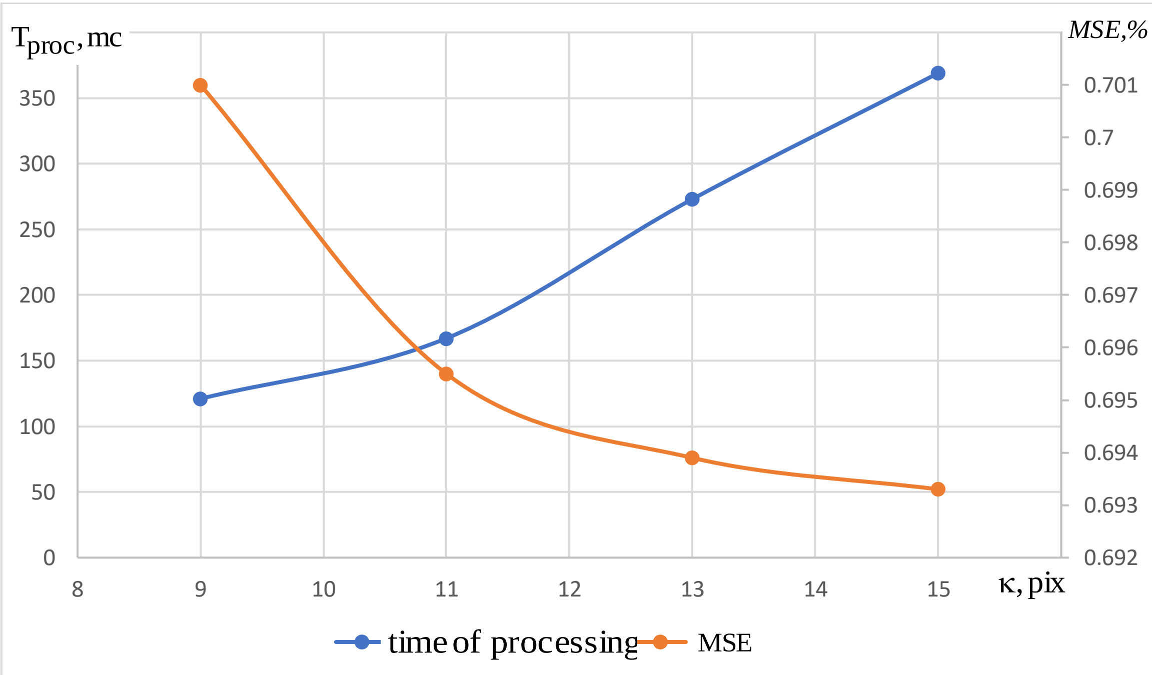

Figure 4: Diagram of the dependence of the standard deviation and the time for data processing on the length of the cube for the developed model of transformation of the alphabet of the video sequence

Thus, the developed model makes it possible to ensure a sufficiently high quality of encoded data ($CKO \approx 0{,}7\,\%$) by transforming the alphabet of the message. The essence of which is to additionally eliminate the psycho-visual redundancy of the information resource data by determining the significance of the elements in the message. The significance of the elements of the message is understood as taking into account both the law of the probabilistic distribution of elements in the message and the structural features of the color model to which the initial data are presented. This makes it possible to reduce the power of the alphabet of the encoded data by almost 10 times, which, in turn, makes it possible to create more favorable conditions for increasing the efficiency of video data encoding from the standpoint of ensuring the required level of quality in conditions of limited bandwidth of the data transmission channel.

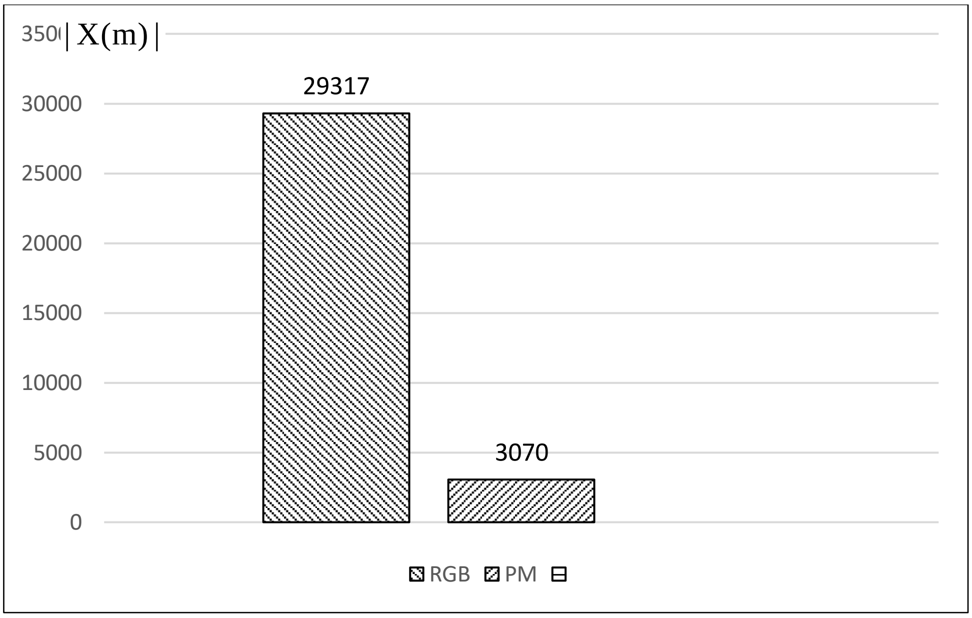

Figure 5: Diagram of the estimation of the power of the alphabet of the encoded data for the RGB model and the developed model

4. References

[1] Wang, S., Zhang, X., Liu, X., Zhang, J., Ma, S. and Gao, W. (2017), Utility Driven Adaptive Preprocessing for Screen Content Video Compression. IEEE Transactions on Multimedia, Vol. 19, No. 3, pp. 660-667.

[2] Gonzales, R.C. and Woods, R.E. (2002), Digital image processing. Prentice Inc. Upper Saddle River, 779 p.

[3] Dong, W. and Wang, J. (2016), JPEG Compression Forensics against Resizing. Computer Science, pp. 1001-1007. DOI: 10.1109/TrustCom.2016.0168. available at: https://www.semanticscholar.org/paper/JPEG-Compression-Forensics-against-Resizing-Dong-Wang/620691ba5ac31d7d26f7167390ce86cb83f7a086 (accessed (accessed 05 June 2021).

[4] Richter, T. (2017), Error Bounds for HDR Image Coding with JPEG XT. Data Compression Conference (DCC), pp. 122-130. DOI: 10.1109/DCC.2017.7.

[5] Xiao, W., Wan, N.A., Hong and Chen, X. (2020), A Fast JPEG Image Compression Algorithm Based on DCT. IEEE International Conference on Smart Cloud (SmartCloud), pp. 106-110. DOI: 10.1109/ SmartCloud49737. 2020.00028.

[6] Rippel, O. (2019), Learned Video Compression. IEEE/CVF International Conference on Computer Vision (ICCV), pp. 3453-3462. DOI: 10.1109/ICCV. 2019.00355.

[7] Bienik, J., Uhrina, M., Kuba, M. and Vaculik, M. (2016), Performance of H.264, H.265, VP8 and VP9 Compression Standards for High Resolutions. 19th International Conference on Network-Based Information Systems (NBiS), pp. 246-252. DOI: 10.1109/NBiS. 2016.70.

[8] Wang, X., Xiao, J., Hu, R. and Wang, Z. (2017), Cruise UAV Video Compression Based on Long-Term Wide-Range Background. Data Compression Conference (DCC), pp. 466-467. DOI: 10.1109/DCC.2017.71.

[9] Minallah, N., Gul, S. and Bokhari, M. (2015), Performance Analysis of H.265/HEVC (High-Efficiency Video Coding) with Reference to Other Codecs. 13th International Conference on Frontiers of Information Technology (FIT), pp. 216-221. DOI: 10.1109/ FIT.2015.46.

[10] Djelouah, A., Campos, J., Schaub-Meyer, S. and Schroers, C. (2019), Neural Inter-Frame Compression for Video Coding. IEEE/CVF International Conference on Computer Vision (ICCV), pp. 6420-6428. DOI: 10.1109/ICCV.2019. 00652.
[11] Narmatha, C., Manimegalai, P. and Manimurugan, S. (2017), A LS-compression scheme for grayscale images using pixel based technique. International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT), pp. 1-5, DOI: 10.1109/ IGEHT.2017.8093980.
[12] Poolakkachalil, T.K., Chandran, S., Muralidharan, R. and Vijayalakshmi, K. (2016), Comparative analysis of lossless compression techniques in efficient DCT-based image compression system based on Laplacian Transparent Composite Model and An Innovative Lossless Compression Method for Discrete-Color Images. 3rd MEC International Conference on Big Data and Smart City (ICBDSC), pp. 1-6, DOI: 10.1109/ICBDSC.2016.7460360.
[13] Wang, Z., Liao, R. and Ye, Y. (2020), Joint Learned and Traditional Video Compression for P Frame. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 560-564. DOI: 10.1109/CVPRW50498.2020.00075.
[14] Bui, V., Chang, L., Li, D., Hsu, L. and Chen, M. (2016), Comparison of lossless video and image compression codecs for medical computed tomography datasets. IEEE International Conference on Big Data (Big Data), pp. 3960-3962. DOI: 10.1109/BigData. 2016.7841075.
[15] Akbari, M., Liang, J., Han, J. and Tu, C. (2020), Learned Variable-Rate Image Compression With Residual Divisive Normalization. IEEE International Conference on Multimedia and Expo (ICME), pp. 1-6. DOI: 10.1109/ICME46284.2020.9102877.
[16] Shinde, T. (2019), Efficient Image Set Compression. IEEE International Conference on Image Processing (ICIP), pp. 3016-3017, DOI: 10.1109/ICIP. 2019.8803230.
[17] Lin, J., Liu, D., Li, H. and Wu, F. (2020), M-LVC: Multiple Frames Prediction for Learned Video Compression. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3543-3551. DOI: 10.1109/CVPR42600. 2020.00360.
[18] Artuğer, F. and Özkaynak, F. (2018), Fractal Image Compression Method for Lossy Data Compression. International Conference on Artificial Intelligence and Data Processing (IDAP), pp. 1-6. DOI: 10.1109/IDAP. 2018.8620735.
[19] Arnob, P., Tanvir, Z., Prajoy, P., Rafi, A., Muktadir Rahman, M. and Mamdudul Haque, Kh. (2015), Iris image compression using wavelets transform coding. 2nd International Conference on Signal Processing and Integrated Networks (SPIN), pp. 544-548, DOI: 10.1109/SPIN.2015.7095407.
[20] Zhu, X., Liu, L., Jin and Na Ai, P. (2016), Morphological component decomposition combined with compressed sensing for image compression. IEEE International Conference on Information and Automation (ICIA), DOI: 10.1109/ICInfA.2016.7832096.
[21] Wang, S., Kim, S. M., Yin, Z., and He, T. (2017), Encode when necessary: Correlated network coding under unreliable wireless links. ACM Transactions on Sensor Networks, Vol. 13(1). DOI: 10.1145/ 3023953.
[22] Phatak, A. (2016), A Non-format Compliant Scalable RSA-based JPEG Encryption Algorithm. International Journal of Image. Graphics and Signal Processing, Vol. 8, No. 6, pp. 64-71. DOI: 10.5815/ijigsp. 2016.06.08.
[23] Barannik, V., Tupytsya, I., Parkhomenko, M., Azatov, A., Pershin, A., Gurzhii, P., Shaikhanova, A. and Karpiinski, M. (2020), The concept of a quantitative sign formation for the internal restructuring of information resource data. Przetwarzanie, transmisja i bezpieczeństwo informacji', pp. 41-52.
[24] Wu, H., Sun, X., Yang, J., Zeng, W. and Wu, F. (2016), Lossless Compression of JPEG Coded Photo Collections. IEEE Transactions on Image Processing, Vol. 25, No. 6, pp. 2684-2696. DOI: 10.1109/ TIP.2016.2551366.
[25] Chen, C. and Zhuo, Y. (2017), A research on anti-jamming method based on compressive sensing for OFDM analogous system. IEEE 17th International Conference on Communication Technology (ICCT), pp. 655-659, DOI: 10.1109/ICCT.2017. 8359718.
[26] Wang, S., Kim, S., Yin, Z. and He, T. (2015), Encode when necessary: Correlated network coding under unreliable wireless links. ACM Transactions on Sensor Networks, Vol. 13, No.1, pp. 24-29, DOI: 10.1145/ 3023953.

[27] Han, S., Mao, H. and Dally, W. (2015), Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding, available at: https://webcache.googleusercontent.com/search?q=cache:DSoAo2aheUoJ:https://arxiv.org/pdf/1510.00149+&cd=2&hl=uk&ct=clnk&gl=ua (accessed 05 June 2021).
[28] Zhurakovskyi, B., Boiko, J., Druzhynin, V., Zeniv, I. and Eromenko, O. (2020), Increasing the efficiency of information transmission in communication channels. Indonesian Journal of Electrical Engineering and Computer Science, Vol. 19(3), pp. 1306-1315. DOI: 10.11591/ijeecs.v19.i3.
[29] Barannik, V., Tupitsya, I., Dodukh, O., Barannik, V. and Parkhomenko, M. (2019), The Method of Clustering Information Resource Data on the Sign of the Number of Series of Units as a Tool to improve the Statistical Coding Efficiency. IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), pp. 32-35. DOI: 10.1109/CADSM.2019.8779243.
[30] Barannik, V., Tupitsya, I., Gurzhii, I., Barannik, V., Sidchenko, S. and Kulitsa, O. (2019), Two-Hierarchical Scheme of Statistical Coding of Information Resource Data with Quantitative Clustering. IEEE International Conference on Advanced Trends in Information Theory (ATIT), pp. 89-92. DOI: 10.1109/ATIT49449.2019.9030451.
[31] Barannik, V., Tupitsya, I., Barannik, V., Shulgin, S., Musienko, A., Kochan, R. and Veselska, O. (2019), The Application of the Internal Restructuring Method of the Information Resource Data According to the Sign of the Number of Series of Units to Improve the Statistical Coding Efficiency. 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), pp. 65-69. DOI: 10.1109/IDAACS.2019.8924460.
[32] Barannik, V., Tupitsya, I., Kovalenko, O., Sidchenko, Y., Yroshenko, V. and Stepanko, O. (2020), The analysis of the internal restructuring method efficiency used for a more compact representation of the encoded data. Advanced Trends in Information Theory (ATIT'2020): proceedings of the Intern. Conf., pp. 48-51. DOI: 10.1109/ATIT50783.2020.9349307.
[33] Yudin, O., Artemov, V., Krasnorutsky, A., Barannik, V. Tupitsya, I. and Pris, G. (2021), Creating a mathematical model for estimating the impact of errors in the process of reconstruction of non-uniform code structures on the quality of recoverable video images. Advanced Trends in Information Theory (ATIT'2021): proceedings of the Intern. Conf., pp. 38-41. DOI: 10.1109/ATIT54053.2021.9678887.
[34] Khmelevskiy, S., Tupitsya, I., Mahdi, Q. A., Musienko O., Parkhomenko, M., & Borovensky , Y. (2021). Development of the external restructuring method to increase the efficiency of information resource data encoding. Information Processing Systems, 3(166), pp. 52-61. DOI: 10.30748/soi.2021.166.06.